DRAFT

Metrics of IT Equipment — Computing and Energy Performance

Bruce Nordman[1]

LBNL High Tech Buildings Project

Lawrence Berkeley National Laboratory

March 10, 2005

This summarizes a full draft paper available as the last research area of:
http://hightech.LBL.gov/datacenters.html
*Comments on the draft and contributions of relevant data are requested by **April 18**;*
these will be incorporated into a final version. Contact the author to be notified of updates.

This report summarizes work to date on the development of a simple, standard method of characterizing the degree to which a single server reduces its energy consumption when operating at low levels of computation compared to what it consumes at peak computing capacity. The goal is to bring more attention and rigor to the issue, and lead to future servers which save energy by having lower power use at lower levels of work load.

Energy efficiency policy and IT industry interest in data center energy consumption has traditionally focused on the maximum power consumption levels of servers and other equipment. More recently, interest has grown in consumption at lower levels of system work load; optimizing these modes offers the potential for energy savings and for more robust system operation and efficient rack space utilization.

A barrier to clear understanding of present consumption and efficiency opportunities is the lack of standard methods to correlate IT equipment energy consumption with the useful information processing tasks being performed. This discussion is an initial effort to remove that barrier. A standard "power vs. load" metric would demonstrate how IT electrical loads vary within the envelope of maximum consumption, and show the potential for, or document the success of, mechanisms to maximize the reduction of energy consumption when IT processing loads are well below a system's maximum capability — the dominant mode for most servers.

Figure 1. Electricity vs. Work Load

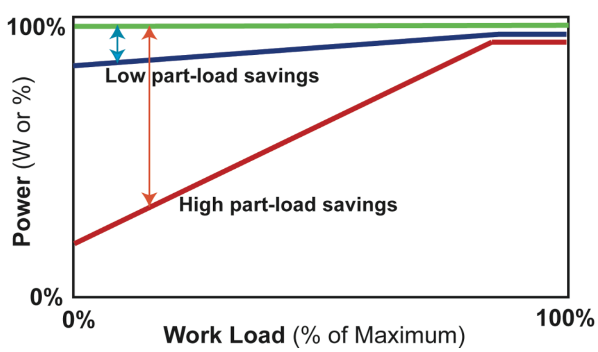


Figure 1 presents such a metric graphically. The work load is a percentage scale, since the metric is not intended for measuring absolute system performance. Examples are web pages served, database transactions completed, or calculations performed. The key is that the computation be driven by external sources so that only a certain amount of work "needs" to be done at any particular time.

Computer benchmarks most commonly compare the speed of one system to another. Measurements that compare a system *only to itself* are inherently simpler to construct, gain acceptance for, and be used. Shifting future IT equipment to have higher savings at low work load times can save large amounts of electricity, even if the maximum consumption values of the system do not change.

Ultimately, a simple power vs. work load metric could be utilized by industry and for energy efficiency interests to document the energy-saving features of products. There is presently no Energy Star specification for servers. The new draft Energy Star Computer Specification includes "desktop-derived" servers and covers idle power and power supply efficiency. The Tier II includes benchmarking systems to "performance per unit energy" (specific procedure not yet specified) and fixing the "network problem". The latter would allow some servers to go to sleep when the latency on waking was within performance bounds.

Servers have a nameplate power that is usually much higher than any load actually ever drawn by the system. They also have several maximum power values which can depend on the type of application used and the particular hardware configuration (e.g. amount of memory installed). There are a continuum of active states from maximum computing down to minimum idle modes of no user application use. Systems may have sleep states but at present these are rarely used.

[1] Bruce Nordman, BNordman@LBL.gov, 510-486-7089

DRAFT

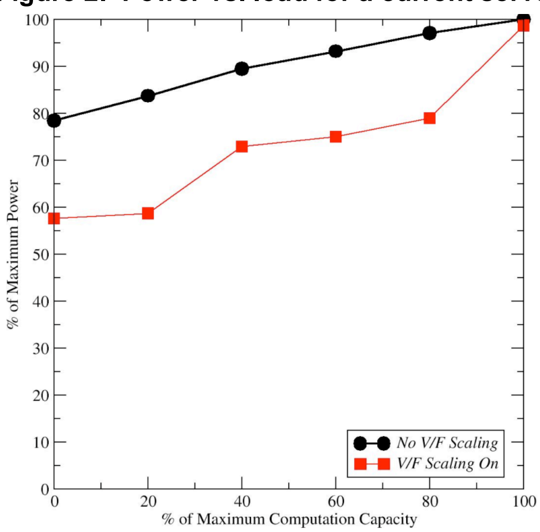
While public data on the degree of server loading are scarce, many web servers operate at around 25% of their observed peak level on average, with most of the rest probably used at no more than 50%. These are relative to the observed peak, *not* the maximum capacity of the system so that average use as a percent of maximum capacity is considerably lower. There are many factors leading to the low average loading of web and similar servers, many quite justifiable for good business reasons.

The most commonly proposed and implemented method of modulating server power consumption in response to demand is to adjust (“scale”) processor frequency and voltage. This offers dramatic power reductions as performance drops. Frequency and voltage scaling were first introduced in processors designed for the mobile market, and have been moving into the desktop and server lines since; the AMD, Intel, and PowerPC lines all include processors that implement it. Processor savings of 75% and system savings of 30% are cited.

Power saving opportunities in clusters (groups) of servers are greater than that available for single-processor systems. One method is to power down (to sleep or off modes) a portion of a cluster when the capacity of all servers is not needed. Since the power levels of products in sleep and off are usually much lower than idle levels, the savings can be significant. In all cases, energy saving strategies need to crafted to not compromise performance (quality of service); a common criterion for this for transaction processing is average response time.

The literature on energy savings obtainable from systems operating at less than their peak capacity is diverse. Savings of a third and higher are often found with real work loads. Another measure is the ratio of idle to maximum power, which has dropped to 60% for some systems (though can be 80% in systems even without voltage and frequency scaling). Figure 2 shows data from one current 4-processor system, from idle to maximum capacity.

Figure 2. Power vs. load for a current server



Many manufacturers have measurements that illustrate the power vs. work load relationship, but are reluctant to publish the raw data. Both axes of the graph can be presented as percentages rather than reported as absolute levels of power or computation. Such graphs need not reveal the precise application being run or system configuration.

Computing benchmarks span a wide range from those that implement sophisticated real-world applications, to those that do simple (but real) work, to those that are “synthetic” and do simple tasks of no utility. Complex benchmarks include those from SPEC and TCP, some of which implement transaction processing and web applications. Benchmarks are used to find the maximum throughput that a system can reliably sustain.

A power vs. load metric could be constructed from any transaction-oriented benchmark. In all cases the system is run full-out to determine what that peak capacity is, then caused to be driven at different “speeds” below that peak.

Conducting the most sophisticated benchmarks requires expensive facilities. It would be highly desirable if energy measurements of systems were routinely made as the performance measures are taken; it would be difficult to justify the effort to repeat these benchmarks solely for energy efficiency interest. Relying only on complex benchmarks for energy efficiency benchmarks would greatly limit how much they could be used.

Simple real or synthetic benchmarks may be good proxies for a complex ones for this purpose. That is, that the shape of the curve produced as in Figure 2 is the same in both cases. The simplest test that correctly reflects system performance should be selected and then used, but confirmation is needed.

Future work should adapt these methods to clusters of servers and to network equipment and storage products. Standard reporting of system configuration and associated power consumption levels is clearly needed to inform data center staff of the real typical and maximum power levels they will encounter with particular hardware. Finally, the topic of comparing the performance (in terms of computational work and energy consumption) of different machines should be explored, though the difficulties in doing this in a general way are daunting.